

# Weekly Report

## 1. 概述

本周我调研了轨迹数据挖掘、可视化方面的文献，继续构思完善 idea。同时，编程调试原来的程序，探究原来 idea 的可行性。并思考为何我上一个 idea 在可行性方面遇到了阻力。

## 2. 详细工作内容

### 2.1 编程调试，完善 idea

我在实现之前的 idea 的途中遇到了困难。目前来说，轨迹数据是最熟悉的数据。为了打开思路，我阅读了其它轨迹数据挖掘的文章。

首先我通过阅读文献[1]，系统地了解了轨迹数据挖掘研究的分类，如图 1。我觉得下一个 idea 可能的思路是：1、找轨迹数据挖掘的一个细分领域，适合引入人类知识（可视分析），或某些数据不够直观，需要可视化（可视化）2、如果可以，引入之前我地图匹配的研究成果 3、原有的数据不适合处理，可以考虑数据变换，把数据转换成另一种表达，如矩阵、张量等。

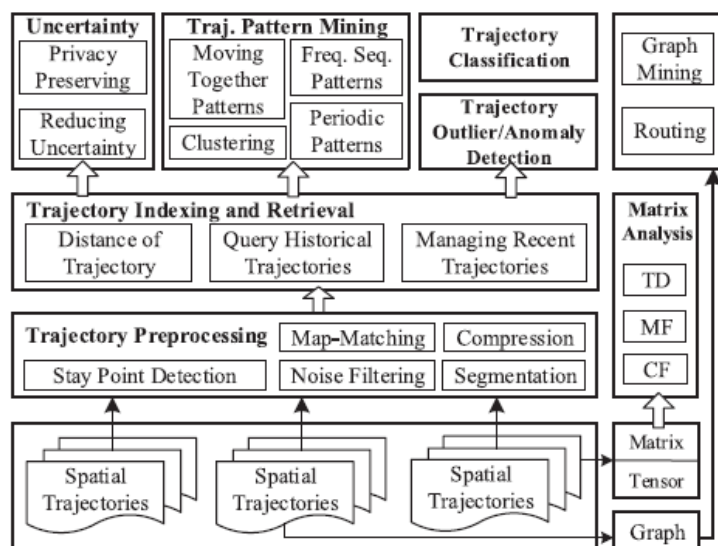


图 1 轨迹数据挖掘研究

### 2.2 通用类的可视化方法研究与应用类可视化方法研究

经过最近组会报告的可视化文章的洗礼。我一直在考虑针对具体应用的可视化文章，和针对某种数据类型的可视化文章的优劣。我认为，一个方法，如果能解决某种抽象数据形式的问题，则必然能够推广到很多具体应用类问题上。因此解决一个抽象问题的文章价值会比解决一个具体问题的文章价值大。尽管如此，我认为目前我还驾驭不了针对抽象问题的文章，优先做出一篇针对应用类的文章。

### 2.3 编程调试

本周我继续调试我之前的地图匹配程序，弄清楚为何我之前的 idea，即用蒙特卡洛法解决图匹配问题的 idea，无法继续进行下去。

首先，我构思 idea 之前，对所使用的技术、要面对的问题把控不足。比如蒙特卡洛估计是什么，怎么用？再比如解决模型泛化性问题目前的研究都有哪些，最适合用可视化方法吗？尽管有用可视化解决迁移学习问题的文章，马博士的文章，但其并未发表在定会上，有足够的参考价值吗？

其次，对子问题的拆解不够合理。我将我的 idea 拆解成了：1) 用可视化评估地图匹配方法的效果 2) 用可视化解决迁移学习的问题。用可视化评估地图匹配的效果真的是必要步骤吗？如果是必要步骤，那么就需要通过实验去验证，以我们现在的条件，能够实验验证吗？不能。另一方面，用可视化解决迁移学习问题，可能就足够解决数据驱动的地图匹配问题的泛化性问题。因为落实到具体使用中，我们能够做到的就是利用可视分析，确定具体需要使用的数据集，或是针对每个数据集选择所需要使用的具体数据。

最后，编程能力差，导致可实现的 idea 也变得不 work。有感于沈汉威老师的报告，编程能力确实是个重要的能力。由于我之前写的程序不够细致严谨，导致处理出来的训练数据出错，进而导致了最终程序准确率不高。同时这个问题的发现也启发了我下一步工作的计划：1) 如果调整完 bug 以后，程序的准确率还是不高，则可以接着我之前 idea 的思路，解决它的泛化性问题；2) 相反，如果证明我的方法合理，就探究是否能将它推广到其它问题，或变成一个通用的方法。

经过这些问题的思考，我认识到论文是个系统性工程，需要全面地考虑问题。

2.4 其它

本周我参加了胡元哲举办的编程训练。

3. 文献阅读

[1] Trajectory Data Mining: An Overview

4. 时间安排

星期	任务	Duration
周一至周日	调试代码、思考工作汇报内容	9:00 - 12:00 和 14:00 - 22:00，共 12 小时

Work Time: more than 50 hours